Proceedings of the Sixth Conference on
African Linguistics. OSU WPL 20.135-147 (1975)

Tone Melodies and the Autosegment*

John Goldsmith
Massachusetts Institute of Technology

1. Introduction

In this paper, I would like to present a theory of tonology--and in fact, a general extension of the theory of phonology--which has as one of its consequences the prediction of the behavior of tones in some ways that are puzzling and even paradoxical from the standard generative view. I shall concentrate here on one kind of behavior that is very familiar to any linguist who has worked on African tone languages: when a vowel desyllabifies or is deleted by some phonological rule, the tone that it was bearing does not disappear--rather, it shifts its location and appears on some other vowel.

Needing a name for this phenomenon, I shall refer to it as 'stability'--that is, the tone melody has its own stability above and beyond that which the vowels provide. Stability, then, in general will be the property of some aspect of the linguistic signal which maintains itself independently of the other aspects. In this case, we note that the tone melody sustains itself independently of modifications occurring to the syllabic structure.

2. Conspiracies

What lesson is to be drawn from this odd behavior? I call it odd, though I have not yet shown that tone stability is an oddity from the standard generative point of view. Let us examine why it is so, and a natural way to do so is to see that ways have been developed within the standard generative framework to deal with tonal phenomena in African tone languages. Only then can we learn what stability tells us about these languages.

I shall consider two examples from the literature, but they could be multiplied easily. In both we find recourse or reference to the notion of 'derivational constraint' or 'conspiracy'. A conspiracy, of course, is a pointed way of referring to a kind of derivational constraint: in this case, we are dealing with derivational constraints of conspiracies to move around the tonal specifications from vowel to vowel in order to find, on the surface, the same tone melody that occurs underlyingly. This is not to say that there are no tonal rules that delete or modify the tonal melody--there surely are such rules. Nonetheless, the normal, 'unmarked' case is where the tone melody survives the effects of phonological rules.

If the tone of a vowel is specified by its features--±High and ±Low, let us say for the sake of discussion--then the pitch of

the vowel is just like any other of its characteristics, like its tenseness, roundedness, and so forth. If a phonological rule should delete that vowel, then its tonal specifications are deleted along with all of its other properties. Suppose we have a phonological rule deleting a vowel as in (1).

(1) V → ∅ / — V V-Deletion

(A common Bantu rule--e.g., see Spa 1973:78, rule (32)). However, we need to save the tonal information of the deleted vowel.[1] Looking at tone as a feature of the vowel, we could do this in one of exactly three ways, all of which are similar in intent.

1. We could posit a special 'Tone copy' rule which copies the tone of the to-be-deleted vowel onto its neighbor--we could do this, I should add, if we permit <u>two</u> tonal feature specifications inside a single vowel segment. Let us permit this for the moment, though I have argued elsewhere that this is a formally incoherent assumption (Goldsmith 1974b); we shall ultimately have no need for such formulations:

(2) Tone Copy $\begin{bmatrix} V \\ \alpha Hi \\ \beta Lo \end{bmatrix}\begin{bmatrix} V \\ \gamma Hi \\ \delta Lo \end{bmatrix} \rightarrow \begin{bmatrix} V \\ \alpha Hi \\ \beta Lo \end{bmatrix}\begin{bmatrix} V \\ \begin{bmatrix} \alpha Hi \\ \beta Lo \end{bmatrix}\begin{bmatrix} \gamma Hi \\ \delta Lo \end{bmatrix} \end{bmatrix}$

A typical derivation applying Tone Copy and V-Deletion would be as in (3).

(3) ...à í...
...à ǐ... Tone Copy
... ǐ... V-Deletion

2. Second, we could have the tone-coyping rule operate after vowel deletion, but make it a global rule looking back in the derivation to the stage just before vowel-deletion. This sounds unnecessarily complicated, and it is; it has been suggested in cases where the V-Deletion is optional and Tone-Copy should be dependent on its application. We shall not pursue this possibility here, for reasons that will be apparent.

3. Third, we could posit a general 'derivational constraint' to apply to <u>all</u> tonal rules--this is the approach Spa takes in his grammar of Enya, a Bantu language. He suggests (I translate): "When a segment carrying a high tone is deleted or becomes incapable of carrying a tone, the high tone is transferred to the nearest syllabic segment... [This constraint] applies each time any rule whatever meets its structural description". (Spa 1973:139 and passim). In fact, the correct statement of his derivational constraint should apply equally to preserve both High and Low tones. This modification both simplifies his phonological system and generalizes his derivational constraint.

Solutions 2 and 3 are explicitly global, and therefore suspect within received generative theory: a theory countenancing global rules approaches vacuity. Solution 3 introduces a general global condition on vowel-affecting rules, and while this seems like an

improvement, in that it is a generalization, it is nonetheless worse theoretically because we now permit not only global rules, but a whole new kind of object which is global and applies anywhere during a derivation outside the set of ordered rules. Only solution 1 holds out a possibility, and yet what we find in actual work is that for every rule of vowel-deletion or desyllabification, we must set up another case of tone-copying, and what we have then is a missing generalization. But the generalization is precisely solution 3, the general derivational constraint. To do satisfactory linguistic work, we need to state a generalization; but inclusion of this generalization within the standard theory amounts to a serious weakening of the theory of phonology. We have reached a crisis in tonological theory.

We might note that even if we did include the derivational constraint, in the belief perhaps that constraining a theory must always play a more minor role than stating a generalization, three important questions would be left unanswered: first, why are the tonal features copied, but not the other features? What makes them special? Second, what is meant by a representation like (4), a contour-toned vowel produced by the derivational constraint--what does it mean for the segment to be segmented? Third, and most telling, the 'conspiracy' to preserve tonal melodies extends past a derivational constraint that whisks the tone off of a sinking vowel: in fact, in a tone language where the Derivational Constraint seems to generally hold, what we find is that vowel assimilation rules like (5) copy all vowel features up to, but <u>not</u> including, tone features.

(4)
$$\begin{bmatrix} +\text{vocalic} \\ \begin{bmatrix} +\text{hi} \\ -\text{lo} \end{bmatrix} \begin{bmatrix} -\text{hi} \\ +\text{lo} \end{bmatrix} \end{bmatrix}$$

(5)
$$V \rightarrow \begin{bmatrix} \alpha\text{tense} \\ \beta\text{back} \\ \gamma\text{high} \\ \delta\text{anterior} \\ \varepsilon\text{round} \\ \vdots \end{bmatrix} / \_\_ \begin{bmatrix} +V \\ \alpha\text{tense} \\ \beta\text{back} \\ \gamma\text{high} \\ \delta\text{anterior} \\ \varepsilon\text{round} \end{bmatrix}$$

A rule like (5) certainly exists in Igbo and Yoruba, and in Enya according to Spa (1973:47, 57). Of course, when two vowels come together, each with its own tone, whether one vowel deletes and its tone gets retained, as in (3), or one assimilates in quality in every regard save tone--the only empirical difference lies in the length and syllable quality of the remaining vowel(s). From the point of view of tone--and its 'conspiracies'--the same fate has come to pass. Yet the derivational constraint speaks only to the case with deletion, not the case of nearly-complete assimilation, as in (5)--thus missing the generalization.

This is the logic of the situation; let us look at some actual cases in more detail. These types of examples could be multiplied.

(A) The first example comes from two articles by Julie Lovins (1971a, b) on Lomongo, whose tonological rules, she suggests, 'conspire, individually or in concert, to derive surface tone patterns on words and phrases without changing the underlying melody." Central to the analysis is what Lovins calls 'tone composition', in which the tones "stay where they are when segmentals are deleted". She continues with an example, "If two vowels are juxtaposed, within a word or across word boundary, it is usual for the first vowel to be elided. Its tone remains and combines with that of the following vowel." For example:

(6)	bàlóngó bǎkáé →	bàlóngǎkáé	'his block'
	bánà bǎmǒ →	bánǎmǒ	'other children'
	bǒmọ̌ bòtámbá →	bǒmọ̌támbá	'another tree'
	bǎtswá là èmí →	bǎtswêmí	'you who lead me away'

With a number of similar examples, Lovins concludes: "The only derived forms that occur are the ones that preserve the underlying melody...and the only way to get these derived forms is to posit a species of rule application that many linguists find objectionable." Lovins is certainly correct, given the standard framework, and she is exceptional among writers on this subject in recognizing the implications for phonological theory of the type of rule she posits.[2]

Lovins gives other examples of the melodic 'conspiracy'; they are discussed in section 5 below as examples of how this matter should in general be treated.

(B) The details of Enya, another Bantu language, that Spa presents are of essentially the same character as those presented by Lovins, at least insofar as they indicate a 'melodic conspiracy' preserving tones against the ravages of vowel deletion. I shall therefore merely mention a few examples of this sort.

The quality of noun-prefix vowels changes before nouns that begin with a syllabic. Whether this is entirely by deletion, as by Rule (1), or by assimilation up to, but not including, tonal information, or both, is not entirely clear.

(7)	/mè + tábè/ →	[mètábè]	'branches'
Class 4	/mè + ẹ́lì/ →	[mẹ̀ẹ́lì]	'moons, months'
	/mè + ìsà/ →	[mìsà]	'poles'
Class 3	/mò + ẹ́lì/ →	[mẹ̀ẹ́lì]	'moon, months'
Class 10	/ǹ + úpú/ →	[m̀m̀pú]	'odor'
	/ǹ + óyá/ →	[ǹǹyá]	'vapor'

In short, deletion and 'total' assimilation of syllabic segments procedes without affecting the tonal contours.

(C) In Yoruba processes of this sort are found, too. The preposition ni is inherently High-toned. When followed by an object starting with a vowel--the usual case--the i elides. If the remaining vowel (that is, the vowel that the object of ni begins with) is any vowel but i, the n of ni becomes l. Thus

(8) /ní + ìlé/ → [nílé] 'at home'
/ní + ònì/ → [lónì] 'at today' (see text)
/ní + òdè/ → [lódè] 'at outside' (see text)

The same processes occur with the verb ni 'have'.

(9) /ní + ọ̀kọ̀/ → [lókọ̀] 'have canoe'
/ní + ọ̀kọ̀/ → [lókọ̀] 'have spear'
/ní + ìrù/ → [nírù] 'have hair'

These tonal markings must be taken with the following interpretation: the vertical accent is Mid; Yoruba has three surface pitches. In all cases the second tone of the object undergoes the effects of the underlying first tone. A High or Mid on the first syllable causes a following Low to be realized as a Falling tone (a flop rule: see (12) below). A Low tone on the first syllable causes a following Mid or High to be downstepped, and a Low tone on the first syllable causes a High on the following syllable to be realized as a Rising tone (again, a flop; this occurs along with the just-mentioned downstep or downdrift).

These processes occur generally in Yoruba; they occur in the object in the ni+object construction just as elsewhere. The surface realization of ni+object has a contracted vowel/tone combination, as we have seen. The result of the contraction is a vowel with the quality of the object's initial vowel (ni+oni → loni, e.g.); the tone of this vowel will be either (a) High, from ni, or (b) the merger of High and the underlying tone. This varies from speaker to speaker. Thus ní+ọ̀kọ̀ → lókọ̀ may be uttered with a Falling tone on the initial o: lôkọ̀, preserving, conspiratorially, the underlying melody.

3. Autosegmental phonology

So much for conspiracies or paradoxical problems for the standard theory of phonology. We proceed to a solution.

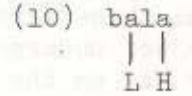
We need a new formalism--one which presents, or re-presents, what we know is central to tone behavior, but one which does not saddle us with formal paradoxes. We know that, tone features aside, the features of the vowels and consonants behave unexceptionally. That is, the standard theory of how vowels and consonants are segmented and how features are unordered inside the segment is precisely correct. We do not want to wander through the wilderness for a theory of consonants and vowels in these respects: we have one that is essentially correct. So we shall start with the premise that the formal representation for vowels and consonants with regard to their non-tonal features is as in, say, *Sound Pattern of English*.[3]

What do we want to do with tonal specifications? One thing we know from the start: the standard answer, which treats the tonal specifications as features of vowels, is inadequate.

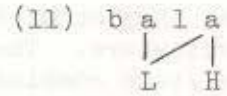
Yet the tone melodies are segmented, in the primary sense of the word: the tone melody is composed of smaller, repeatable units,

the tonal segments. The obvious proposal to make, then is that the tone melody also composes a segmented line--a second tier of segments. To complete the representation, we will need association lines between the string of phonological segments, and the string of tonological segments. For a vowel to be associated with a particular tone will mean that it is uttered at that tone's pitch.

For example, now, if we have a two-syllable word 'bàlá', this is represented:

(10)


If it should undergo a rule (we might call this a 'flop' rule for geometric reasons) and gain a rising tone on the second syllable by 'assimilation', then we will represent this as in (11), and the Flop rule that creates it will be as in (12).

(11)


(12)


The dotted line in (12) represents the structural change. In (11) there is no one-to-one correspondence between tonemes and syllables, but that is exactly what we know happens in tone languages--this remarkable feature of the formalism is precisely what we would desire in a formalism.

Now we shall see that the formalism that is sketched above--which I call 'autosegmental phonology'--is the solution to the paradoxical situation discussed in sections 1 and 2. The existence of the tone melody's 'stability' was our concern: how could it be that a tone refused to be deleted when its vowel was deleted?

But in our new formalism, this is precisely what is predicted. In any theory of generative phonology (and this one is no different), a deletion rule deletes a segment. Now, if a rule ((1), for example) should delete a vowel, it does not delete any of the tone segments that the vowel is associated with, since those tone segments are quite separate segments. The worst that can happen is that the tone segments will be left 'orphaned' or free, without a vowel associated with it. And this will be the interesting case to look at in detail.

I would like to emphasize the point we have just seen: the 'stability' phenomenon, formerly paradoxical, has become a natural consequence of an autosegmental generative system--not by proposing a constraint on rules, but rather by proposing in effect a new geometrical shape (in a somewhat abstract sense) for formal representations. We may observe, too, that this representation could be likened to a theory of co-articulation of the mouth and some of the laryngeal (pitch) features.

One further extremely important point about autosegmental representation must be made. A Well-formedness Condition is placed on representations which expresses a good deal about their 'geometry'.

Well-formedness Condition

(1) All tones must be associated with at least one syllabic segment in the other tier;
Conversely, all syllabic segments in the upper tier must be associated with at least one toneme in the tone tier.

(2) No association lines may cross.

This well-formedness condition does not throw out representations that do not satisfy it ((13), e.g.); rather, we interpret the well-formedness condition to add or delete association lines in a minimal way (generally by minimizing the number of line additions or deletions) so as to meet the well-formedness condition as completely as possible. Thus (13) is converted by the well-formedness condition to (14) by the addition of association lines. This well-formedness condition is not a rule, nor a derivational constraint; it is a definition of the formalism--which is not to say that it does not have striking and immediate empirical consequences, as we shall see.

(13) ma gi ri
H

(14) ma gi ri
H

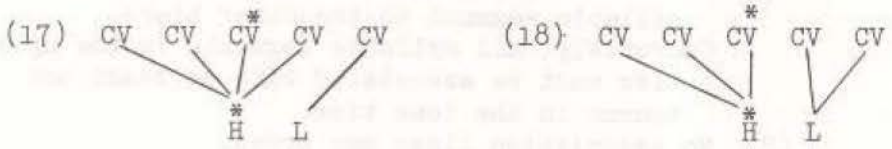
In originally associating a syllable with its tone, languages have recourse to two types of procedures (see Goldsmith 1975): either they spread out one-to-one from a boundary, as in Mende (see Leben 1973) (figure (15)) or a tone is specifically associated with a correspondingly marked syllable.

(15) CV CV CV → CV CV CV → CV CV CV
L H L H L H
by major association rule
by the well-formedness condition

That is, a syllable of a word can be marked as the specific one that a certain tone is to be coordinated with. I shall use a 'star' in either tier to indicate which segments are marked to associate with each other.

(16) CV CV CV CV CV
H L
(as in English, for example)

In Bantu languages, it is not uncommon that a specific syllable is marked to be the one which the tone connects to. In such cases, as in (16), the well-formedness condition intends to spread the tone melody over the yet-unassociated syllables; however, there is an apparent ambiguity regarding the final two syllables: does the H spread there (as in (17)) or the L (as in (18))?



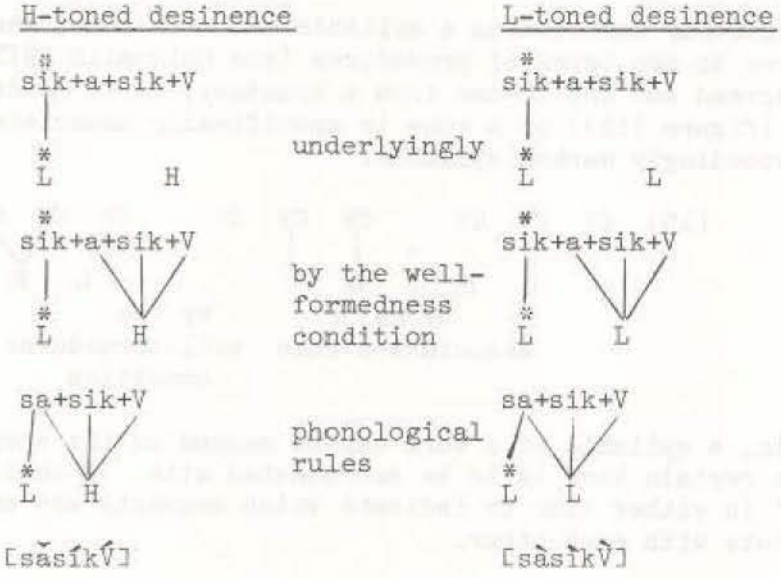
Without exception (English, Sanskrit, Japanese, and Tonga have been examined in detail in this regard), the unstarred toneme spreads; thus (18) is produced by the well-formedness condition, not (17). Functionally, we may regard this as a preservation of the prominence of the accented syllable, which would be lost if (17) were derived.

With this machinery at our disposal, we may return to Lomongo.

4. Lomongo reduplication: a reanalysis

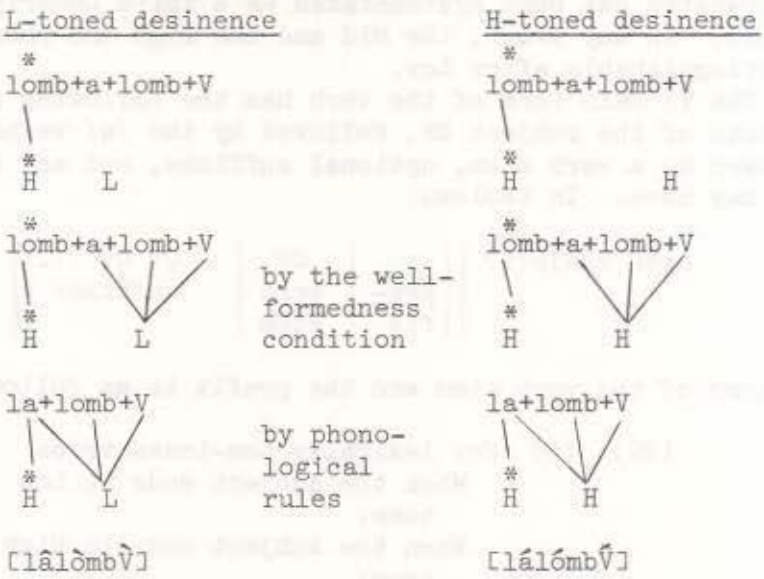
Let us consider in more detail the reduplication treated by Lovins (1971a, b). Verbs are lexically marked for tone, H or L: the stem is reduplicated and an a infix is added between the two copies of the stem. An L or H tone desinence then follows.

(19) L-toned stem /sɪ̀k/ 'stop'



(The last stage is reached by purely phonological rules: k → Ø; and ia → a).

(20) /lómb/ H-toned stem 'be shy'



In short, from the notation we get the 'conspiratorial' results automatically by keeping the syllabic and the tonal levels formally separate. It may be noted that we get the desinence-tone spreading automatically, too, as well ay a formal understanding of the notion 'contour tone'. Furthermore, the process of total vowel assimilation--construed, as, e.g., (21)--has the desired property of copying all features up to, but not including, tone features, since tone features aren't features of vowels.

(21) X - [+syllabic] - [+syllabic] - Y

1 2 3 4 →
1 3 3 4

5. An autosegmental analysis of Igbo

In section 4, we observed that conspiracies of the sort discussed in section 2 are straightforward consequences of the autosegmental representation. In deleting a vowel, we do not delete its tone; rather, its tonal segment finds some other place to associate.

This conclusion came about by viewing tones as segments on an equal rank with 'phonological' segments. This parallelism can be pursued; in fact, we find in general perfect formal symmetry between the two levels. The 'dual', then, of vowel deletion would be tone-deletion, followed by reassociation to another tone by the vowel that had been associated with the deleted tone.

A particularly instructive example of this sort is found in Uhũhũ Ibgo (Green and Igwe 1963). Igbo has three underlying tones--High, Mid, and Low, where Mid in this dialect is a downstepping High. That is, when Mid is uttered after a non-Low tone (High or Mid), it has the pitch of a High that has been downstepped as if

by an intervening 'phantom' Low tone. In this dialect, however, the downstep has been systematized as a third underlying tone, the Mid. In any event, the Mid and the High are phonetically indistinguishable after Low.

The II Main form of the verb has the following pattern: it consists of the subject NP, followed by the /a/ verbal prefix, followed by a verb stem, optional suffixes, and any objects the very may have. In tableau:

(NP subject)	a prefix	CV verb stem	CV CV ... suffixes	(Direct Object)	Etc.

The tone of the verb stem and the prefix is as follows:

(22) (i) For lexically Low-toned verbs:

	AFFIX	STEM
When the subject ends in Low tone:	"High"	Low
When the subject ends in High tone:	Low	Low

(ii) For lexically Non-Low verbs:

When the subject ends in Low tone:	Low	"High"
When the subject ends in High tone:	Mid	"High"--that is, on the same pitch as affix

I have put "High" in quotation marks where it occurs in a position where High and Mid are indistinguishable on the surface. Furthermore, we must note that a High following a Mid is realized on the same pitch as the Mid.

Such a superficial description is misleading. The alternations displayed there are quite simple when stated autosegmentally. Note first that the second row in both (i) and (ii) is simpler; in particular, the affix and the stem are on the same pitch for both Low and Non-Low verb stems. (It is crucially important to recognize that the High-toned form in (ii) could <u>not</u> be written 'M M', because that would mean a downstep between the affix and the stem, which does <u>not</u> occur.)

What is happening here is that the underlying forms are as in (23), which is realized on the surface when the subject ends in a L-tone.

(23)

<u>Low-toned verbs</u>		<u>Non-Low toned verbs</u>	
affix	verb	affix	verb
H	L	L	M

The Low-toned verb shows up as such; the non-Low toned verb shows up as Mid, a common occurrence throughout the Igbo verb forms. The affix is, in effect, dissimilated from the stem. To derive the tonal forms when the subject ends in a High-tone, there is a rule (24) which deletes the affix tone following a subject ending in H:

(24) H [[T] affix] verb — Condition: Last-cycle: I Main form occurs only in non-embedded forms.

1 2 →
1 Ø

For example,

(25) Low-Toned Verb

(a) # Adha # a+za+a # 'Adha has swept...'
L H # (H) L H # Underlying
Ø Rule (24)

Adha # a+za+a
L H # L H

Adha # a+za+a # By the well-formedness
L H # L H # condition.
'Àdhá àzàá...'

(b) High-toned Verb

'Adha has carried bags' 'The chief has carried bags'

Adha # a+ci+a # akpa # # Eze # a+ci+a # akpa
L H # (L) M H # L L # # H L # L M H # L L
Ø by rule (24)

Adha # a+ci+a # akpa
L H # M H # L L # No change

Adha # a+ci+a # akpa
L H # M H # L L # No change

[Àdhá ácíá àkpà] [Ézè àcíá àkpà]

Notice that the final form in (25b) has an M associated with two vowels. This can mean only one thing: the two syllables are uttered

at the same pitch, of course, with the 'a' downstepped from the subject-final H of 'Adha'. And this is the correct result as indicated in (22).

6. Conclusion

It is not uncommon to find the observation made that in African tone languages, the tone melody seems to act almost independently of the syllable structure. It is another matter to incorporate this observation into a specific formal theory. The autosegmental framework is such a proposal; and this paper sketches the way it deals with some representative tonal rules.

Footnotes

*This work was supported by a grant from the National Institutes of Health, 5T01 HD00111-10.

[1]Note what Carl Meinhof wrote in *Grundriss einer Lautlehre der Bantusprachen*: "A syllable may disappear without leaving a trace, and yet its former existence there may still be indicated by its tone, which has been imparted to the next syllable. In this way it is sometimes found that the tones are the most persistent element in a language."

[2]We may note that one of the major rules Lovins posits--the monotony rule, spreading the desinential tone over the post-radical syllables--is no exception to the melody conspiracy, as she suggests it is. It is rather due to the well-formedness condition on tonological representations; see (19) and (20).

[3]Of course, there are small points to question, such as what features there are, whether features are binary, whether there are linkages between features, whether there are unspecified features, and so on. But these are, in context, small issues.

References

Chomsky, N., and M. Halle. 1968. *The Sound Pattern of English* New York: Harper & Row.

Cohen, Michael. 1974. Tone in Tonga nouns. Unpublished paper, MIT.

Goldsmith, John. 1974a. English as a tone language. Unpublished paper, MIT.

_____. 1974b. An autosegmental tonology of a fragment of Igbo. Unpublished paper, MIT.

_____. 1975. An autosegmental typology of tone. *Papers from the Fifth North Eastern Linguistics Society*. Harvard Univ.

Green, M. M. and G. E. Igwe. 1963. *A Descriptive Grammar of Igbo*. Oxford University Press.

Leben, Will. 1973. *Suprasegmental Phonology*. Unpublished Ph.D. dissertation, MIT.

Lovins, Julie. 1971a. Melodic conspiracies in Lomongo tonology. *Papers from the Seventh Regional Meeting, Chicago Linguistic Society*. Chicago: Chicago Linguistic Society.

_____. 1971b. The tonology of Lomongo reduplication. *Studies in African Linguistics* 2.3:257-270.

May, R. and J. Goldsmith. To appear. Tone sandhi in Sanskrit.

Rowlands, E. C. 1969. *Yoruba*. English Universities Press.

Spa, Jaap. 1973. *Traits et Tons en Enya, Phonologie Générative d'une Langue Bantoue*. Tervuren, Belgique:Musée Royal de L'Afrique Centrale.